What is claimed is:

1. A method for estimating the copy number of a genomic region in an experimental sample comprising:

(a) isolating nucleic acid from the experimental sample;

(b) amplifying at least some regions of the nucleic acid;

(c) labeling the amplified products;

(d) hybridizing the labeled amplified products to an array to obtain a hybridization pattern, wherein the array comprises a plurality of genotyping probe sets for a plurality of SNPs, wherein a probe set comprises:

(i) a plurality of perfect match probes to a first allele of a SNP,

(ii) a plurality of perfect match probes to a second allele of the SNP,

(iii) a plurality of mismatch probes to the first allele of the SNP, and

(iv) a plurality of mismatch probes to the second allele of the SNP,

(e) obtaining a measurement for the SNP in the experimental sample wherein the measurement, *S*, is the log of the arithmetic average of the intensities of at least two of the perfect match probes for the SNP in the hybridization pattern;

(f) obtaining an *S* value for the SNP in each of a plurality of reference samples that are matched to the experimental sample in genotype call;

(g) calculating the mean and the standard deviation for the reference sample *S* values using the values obtained in (f);

(h) obtaining a log intensity difference by subtracting the mean value obtained in (g) from the value obtained in (e); and

(i) estimating the copy number of the region including the SNP assuming a linear relationship between log intensity ratio and log copy number.

2. The method of claim 1 wherin the *S* values for all SNPs genotyped in the experimental sample and in each reference sample are normalized so that the mean for all the autosomal SNPs in a sample is zero and the variance is 1.

3. The method of claim 1 further comprising calculating a p-value for the estimated copy number alteration and determining if the p-value is less than a threshold p-value, wherein the estimated direction of copy number change is significant if the p-value is less than the threshold.

4. The method of claim 2 further comprising calculating a p-value for the estimated copy number alteration and determining if the p-value is less than a threshold p-value, wherein the estimated direction of copy number change is significant if the p-value is less than the threshold.

5. The method of claim 1 wherein the *S* value is calculated using:

$S = Log(\frac{1}{X}\sum_{i=1}^{X} PM_i)$ where PM_i is the intensity of the perfect match cell of probe pair *i* and X is the number of perfect match probes in a set.

6. The method of claim 5 wherein X is between 1 and 30.

7. The method of claim 5 wherein X is 20.

8. The method of claim 1 wherein copy number is estimated using:

$Copy\ Number \approx \exp(b + m \times (\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}))$ wherein $\widetilde{S}_{jg}^{C}$ is the log of the average of the intensities of the perfect match probes for a SNP j of genotype g in an experimental sample c, normalized to the S values of all SNPs genotyped in the experimental sample, $\hat{\mu}_{jg}$ is the average mean of the normalized S values for SNP j in a plurality of reference samples of genotype g at SNP j, b is the y-intercept and m is the slope of a line defined by plotting intensity values from SNPs of known copy number.

9. The method of claim 8 further comprising the step of calculating a p-value for the direction of estimated copy number alteration using:

$p_j = \min(1 - \Phi(\frac{\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}), \Phi(\frac{\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}))$ and determining if p_j is equal to or less than a threshold p-value.

10. The method of claim 8 wherein b is equal to about 0.693 and m is equal to about 0.895.

11. The method of claim 10 further comprising the step of calculating a p-value for the direction of estimated copy number alteration using:

$p_j = \min(1 - \Phi(\frac{\tilde{S}_{jg}^C - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}), \Phi(\frac{\tilde{S}_{jg}^C - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}))$ and determining if p_j is equal to or less than a threshold p-value.

12. The methof of claim 1 wherein the experimental sample is a tumor sample.

13. The method of claim 1 wherein the experimental sample is a mixture of tumor and normal cells.

14. The method of claim 1 wherein the experimental sample is a sample that is from a non-cancerous sample.

15. The method of claim 1 wherein the experimental sample is a sample that is suspected of having a chromosomal anomoly selected from the group consisting of a constitutional anomoly, an acquired anomoly, a numerical anomoly, a structural anomoly and mosaicism.

16. The method of claim 8 wherein at least some of the SNPs of known copy number are SNPs on the X chromosome.

17. The method of claim 1 wherein each *S* value obtained in (f) that is more than 3 standard deviations from the mean of the *S* values is excluded from the estimation of mean and standard deviation of the reference distribution calculated in (g).

18. The method of claim 1 wherein a second estimate of copy number is obtained by comparing the discrimination ratio, DR, of a SNP in an experimental sample with an average DR from that SNP in a plurality of genotype matched reference samples, where the DR for a probe set with 20 PM/MM probe pairs is calculated using:

$$DR = \frac{1}{20}\sum_{i=1}^{20}\left(\frac{PM_i - MM_i}{PM_i + MM_i}\right)$$

19. A method of identifying a genomic region that is amplified or deleted in an experimental sample comprising:

hybridizing a nucleic acid sample derived from the experimental sample to a genotyping array and measuring hybridization intensities for a plurality of perfect match probes, PM_i;

calculating a value, S, for each SNP genotyped by the array using:

$S = Log\left(\frac{1}{X}\sum_{i=1}^{X} PM_i\right)$ where X is the number of PM probes for an individual SNP;

normalizing a plurality of S values so that the mean of the S values is zero and the variance is one;

obtaining normalized mean S values for each SNP genotyped by the array in a plurality of reference samples;

estimating copy number of at least one SNP in the experimental sample;

determining the direction of change for the SNP in the experimental sample; and

measuring a p-value to determine confidence level in the predicted direction of change.

20. The method of claim 19 wherein copy number is estimated by assuming a linear relationship between the log estimated copy number and the log intensity ratio.

21. The method of claim 19 wherein copy number is estimated using:

$Copy\ Number \approx \exp(b + m \times (\tilde{S}_{jg}^{C} - \hat{\mu}_{jg}))$ where b is about 0.693 and m is about 0.895.

22. The method of claim 19 wherein the nucleic acid sample is derived from the experimental sample using the whole genome sampling assay (WGSA).

23. A method for determining if the copy number estimates of two or more consecutive SNPs is significant comprising:

identifying two or more contiguous SNPs that either all show an estimated reduction in copy number or all show an estimated increase in copy number relative to a plurality of reference samples;

calculating $\tilde{z}_{m,n}$ using $\tilde{z}_{m,n} = \frac{1}{\sqrt{n-m+1}} \sum_{j=m}^{n} \hat{z}_{jg} \sim N(0,1)$;

converting $\tilde{z}_{m,n}$ to a probability using the standard Φ function to obtain a p-value; and,

concluding that the estimates are significant using a p-value threshold.

24. A method of identifying at least one region of loss of heterozygosity comprising:

identifying at least one contiguous stretch of homozygous SNP genotype calls in the genome of an experimental sample;

obtaining a probability, $\hat{P}_i$ of homozygosity for each SNP in the contiguous stretch wherein $\hat{P}_i = \frac{\#of\ AA\ or\ BB\ calls\ on\ SNP\ i}{total\ \#of\ genotype\ calls\ on\ SNP\ i}$;

calculating the probability that each of the SNPs in the contiguous stretch is homozygous by using: $\hat{P}$ (SNP m to n homozygous)=$\prod_{i=m}^{n} \hat{P}_i$; and,

identifying the region containing the SNPs as a region of loss of heterozygosity if $\hat{P}$ (SNP m to n homozygous) is less than a p-value threshold.

25. The method of claim 24 wherein the contiguous stretch is at least 10 SNPs that are genotyped.

26. A method for estimating the copy number of a region identified as a region of loss of heterozygosity by the method of claim 24 comprising:

calculating an *S* value for at least one of the SNPs in the identified region in the experimental sample using: $S = Log(\frac{1}{X}\sum_{i=1}^{X} PM_i)$ where PM_i is the intensity of the perfect match cell of probe pair *i* and X is the number of probe pairs in a set and normalizing the *S* value;

calculating normalized S values for the at least one SNP from a plurality of matched genotpye call reference samples and calculating an average of the reference sample normalized S values for the SNP;

comparing the normalized S value for the SNP in the experimental sample with the average of the normalized S values for the SNP in the reference sample to obtain a ratio; and

estimating copy number of the SNP in the experimental sample.

27. The method of claim 26 wherein copy number is estimated for 2 or more contiguous SNPs in the region.

28. The method of claim 26 wherein a p-value is calculated for the copy number estimate using $p_j = \min(1 - \Phi(\frac{\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}), \Phi(\frac{\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}))$.

29. The method of claim 26 wherein the plurality of matched genotype reference samples comprises at least 10 samples.

30. A computer software product comprising:

computer program code for inputting a plurality of perfect match intensity values (PM_i) for a plurality of SNPs in an experimental or a reference sample;

computer code for calculating the log of the mean of the intensity values for each individual SNP in each sample, wherein there is a plurality of reference samples;

computer code for normalizing mean values within individual experimental and reference samples;

computer program code for calculating a log of the mean of the intensity value for each individual SNP in all reference samples of matched genotype call at that individual SNP;

computer program code for calculating a log intensity difference between the log mean intensity of a SNP from an experimental sample and the log mean intensity of that SNP from reference samples matched to the experimental sample in genotype call at the SNP;

computer program code for estimating the copy number of the SNP using a log-log linear model;

computer program code for calculating a p-value for the direction of change indicated by the estimated copy number;

computer program code for determining if the calculated p-value is less than a selected threshold value; and

a computer readable media for storing said computer program codes.

31. The computer software product of claim 30 wherein the log of the mean intensity value for each SNP is calculated using $S = Log(\frac{1}{X}\sum_{i=1}^{X} PM_i)$ where X is the number of PM probes per SNP.

32. The computer software product of claim 30 wherein p valued is calculated using: $p_j = \min(1-\Phi(\frac{\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}), \Phi(\frac{\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}))$.

33. The computer software product of claim 30 wherein copy number is estimated using: $Copy\ \ Number \approx \exp(b + m \times (\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}))$.

34. A computer software product for identifying at least one region of loss of heterozygosity comprising:

computer program code for identifying at least one contiguous stretch of homozygous SNP genotype calls in the genome of an experimental sample;

computer program code for obtaining a probability, $\hat{P}_i$ of homozygosity for each SNP in the contiguous stretch wherein $\hat{P}_i = \frac{\#\ of\ AA\ or\ BB\ calls\ on\ SNP\ i}{total\ \#\ of\ genotype\ calls\ on\ SNP\ i}$;

computer program code for calculating the probability that each of the SNPs in the contiguous stretch is homozygous by using: $\hat{P}$ (SNP m to n homozygous)$= \prod_{i=m}^{n} \hat{P}_i$;

computer program code for identifying the region containing the SNPs as a region of loss of heterozygosity if $\hat{P}$ (SNP m to n homozygous) is less than a p-value threshold; and

a computer readable media for storing said computer program codes.

35. A system for estimating copy number in an experimental biological sample comprising:

a processor; and a memory being coupled to the processor, the memory storing a plurality of machine instructions that cause the processor to perform a plurality of logical steps when implemented by the processor, said logical steps comprising;

calculating the log of the mean of the intensity values of a plurality of perfect match intensity values (PM_i) for a plurality of SNPs in an experimental or a reference sample for each individual SNP in each sample, wherein there is a plurality of reference samples;

normalizing mean values within individual experimental and reference samples;

calculating a log of the mean of the intensity value for each individual SNP in all reference samples of matched genotype call at that individual SNP;

calculating a log intensity difference between the log mean intensity of a SNP from an experimental sample and the log mean intensity of that SNP from reference samples matched to the experimental sample in genotype call at the SNP;

estimating the copy number of the SNP using a log-log linear model;

calculating a p-value for the direction of change indicated by the estimated copy number; and,

indicating if the calculated p-value is less than a selected threshold value.

36. The system of claim 35 wherein the log of the mean intensity value for each SNP is calculated using $S = Log(\frac{1}{X}\sum_{i=1}^{X} PM_i)$ where X is the number of PM probes per SNP.

37. The system of claim 35 wherein p valued is calculated using:

$$p_j = \min(1 - \Phi(\frac{\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}}), \Phi(\frac{\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg}}{\hat{\sigma}_{jg}})).$$

38. The system of claim 35 wherein copy number is estimated using:

$$Copy\ Number \approx \exp(b + m \times (\widetilde{S}_{jg}^{C} - \hat{\mu}_{jg})).$$

39. The system of claim 38 wherein b is about 0.693 and m is about 0.895.